
education policy analysis archives

A peer-reviewed, independent,
open access, multilingual journal



Arizona State University

Volume 25 Number 65

June 26, 2017

ISSN 1068-2341

An Analysis of Predictors of History Content Knowledge: Implications for Policy and Practice

Paul G. Fitchett

Tina L. Heafner



Richard G. Lambert

University of North Carolina at Charlotte
United States

Citation: Fitchett, P. G., Heafner, T. L., & Lambert, R. G. (2017). An analysis of predictors of history content knowledge: Implications for policy and practice. *Education Policy Analysis Archives*, 25(65). <http://dx.doi.org/10.14507/epaa.25.2761>

Abstract: How and to what extent students learn history content is a complicated process, drawing from the instructional opportunities they experience; the policy prioritization of history/social studies instruction in schools; and their own cultural perspectives toward the past. In an attempt to better understand the complex inter-play among these dimensions, we examined relationships among student sociocultural characteristics, instructional exposure, and school-level variables and US History content knowledge. Using data from the 2010 National Assessment of Educational Progress Test on US History (NAEP-USH), multilevel analyses indicated that while sociocultural indicators (such as race, gender, and socioeconomic status) correlate with achievement, students'

instructional exposure variables remain significant predictors of history content knowledge. Moreover, school context such as building-level demographics and state testing-policy predict between school variance in content knowledge and moderate the achievement gap. Results also suggest that, while a substantial achievement gap remains, exposure to text-based instructional practices is associated with increased knowledge. Findings from this study have policy implications for the development of a more inclusive social studies curriculum, the advocating of text-dependent instruction as a high-leverage practice among history teachers, and cautious consideration of tests as proxies for accountability in history education.

Keywords: history education; social studies education; content knowledge; National Assessment of Educational Progress (NAEP)

Un análisis de los predictores del conocimiento del contenido de la historia: Implicaciones para la política y la práctica

Resumen: Entender cómo y cuántos alumnos aprenden sobre contenido histórico es un proceso complicado, basado en oportunidades de aprendizaje que los alumnos encuentran, en una política de priorización de la enseñanza de contenido histórico o de ciencias sociales, y en las perspectivas culturales de los alumnos sobre el estudio del pasado. Con el objetivo de entender la relación compleja entre los diversos aspectos en este proceso de aprendizaje, investigamos la relación entre características socioculturales de los alumnos, acceso a este tipo de enseñanza, y variables explicativas de escuelas y del conocimiento de la historia de los Estados Unidos. El resultado de nuestra modelación jerárquica indica que mientras variables socioculturales (raza, sexo y situación socioeconómica) se correlacionan con éxito, con datos de la Evaluación Nacional de la Prueba de Progreso Educativo en Historia Americana (NAEP-US, por sus siglas en inglés) Académico, las variables de acceso a este tipo de enseñanza son parámetros importantes para el conocimiento del contenido histórico. Además, variables sobre la escuela, por ejemplo demografía de la escuela y políticas estatales en cuanto a pruebas académicas, son parámetros relevantes en el análisis de varianza en conocimiento sobre el contenido académico y la disparidad en éxito académico entre alumnos. Los resultados también indican que a pesar de la existencia de la disparidad en el éxito académico, el acceso a textos y materiales educativos está relacionado con el avance del conocimiento. Nuestras conclusiones son relevantes para el desarrollo de políticas educativas de un currículo en ciencias sociales que sea más inclusivo, el apoyo al uso de instrucciones usando textos / libros académicos, y la cautela en cuanto al uso de pruebas para medir la enseñanza de la historia.

Palabras-clave: enseñanza de historia, enseñanza de ciencias sociales, conocimiento de contenido académico, Evaluación Nacional de la Prueba de Progreso Educativo en Historia Americana (NAEP-US)

Uma análise dos preditores do conhecimento do conteúdo da história: Implicações para política e prática

Resumo: Entender como e o quanto alunos aprendem sobre conteúdo histórico é um processo complicado, baseado em oportunidades de aprendizagem que alunos encontram, numa política de priorização do ensino de conteúdo histórico ou de ciências sociais, e nas perspectivas culturais dos alunos sobre o estudo do passado. Com o objetivo de entender a relação complexa entre os vários aspectos neste processo de aprendizagem, nós investigamos a relação entre características socioculturais dos alunos, acesso a este tipo de ensino, e variáveis explanatórias de escolas e do conhecimento da história dos EUA.

Usando dados da Avaliação Nacional do Teste de Progresso Educacional em História Americana (NAEP-US, por sua sigla em inglês) de 2010, os resultados da nossa modelação hierárquica indicam que enquanto variáveis socioculturais (raça, sexo e situação socioeconômica) são correlacionadas com sucesso acadêmico, variáveis de acesso a este tipo de ensino são parâmetros importantes para o conhecimento de conteúdo histórico. Além disso, variáveis sobre a escola, por exemplo demografia da escola e políticas estaduais quanto a testes acadêmicos, são parâmetros relevantes na análise de variância em conhecimento sobre o conteúdo acadêmico e a disparidade em sucesso acadêmico entre alunos. Os resultados também indicam que apesar da existência da disparidade em sucesso acadêmico, acesso a textos e materiais educativos é relacionado com o avanço de conhecimento. Nossas conclusões são relevantes para o desenvolvimento de políticas educacionais de um currículo em ciências sociais que seja mais inclusivo, o apoio ao uso de instruções usando textos/ livros acadêmicos, e a cautela quanto ao uso de provas para medir o ensino de história.

Palavras-chave: ensino de história, ensino de ciências sociais, conhecimento de conteúdo acadêmico, Avaliação Nacional do Teste de Progresso Educacional em História Americana (NAEP-US)

Introduction

Teaching history is a complex and multilayered process balancing priorities among content knowledge, curriculum requirements, assessments, cognitive demands, and learning approaches. Among these competing interests is an intricate relationship between how students learn history and what historical content students are taught in school contexts. Within the myriad of factors associated with teaching and learning in history education (Levistik, 2008; VanSleight, 2008), how students are taught, referred to as instructional exposure, affects students' ability to make sense of history. Previous research highlights the importance of inquiry, discipline-specific writing, content literacy, and critical thinking as essential to historical understanding (De La Paz, 2005; Nokes, 2010; Reisman, 2012; Shanahan, 2009; Wineburg & Reisman, 2015). Further complicating students' instructional exposures to history are the various curriculum and content mandates imposed through state policies. Unintended consequences of accountability pressures, such as teaching to the test, can have negative consequences for teaching and learning (Au, 2007; Nichols, Glass, & Berliner, 2012). Alternative analyses of National Assessment for Education Progress (NAEP) and other assessment data link accountability standards and testing to positive student achievement outcomes in other subject areas (Braun, 2004; Dee & Jacob, 2011; Hanushek & Raymond, 2006). However, empirical research examining the effects of testing and accountability on history achievement is largely absent.

History content knowledge is also influenced by the sociocultural context of the learner. Students' identity (ethnicity, class, gender, race, and etc.) affects how students selectively make meaning of content (VanSleight, 2011; Wineburg, Mosborg, Porat, & Duncan, 2007). In other words, learners' levels of engagement with historical content and pedagogy are contingent upon their perceived positionality within the history curriculum as well as their own experiences. School context further confounds the learning process. School-level effects are well-documented predictors of student learning across disciplines (Fantuzzo, Lebeof, & Rouse, 2014; Goldsmith, 2011). School demographic factors correlate with teacher decision-making and student learning outcomes in social studies (Epstein, 2009; Fitchett, Heafner, & Lambert, 2014a, 2014b; Levinson, 2012; Segall, 2006).

In total, research suggests that students' historical learning is an intricate puzzle, contingent upon numerous within-school and outside-of-school dimensions. However, few studies have looked at these puzzle pieces collectively in order to uncover a more elaborate picture of their relationship

to history content knowledge. Using data from the grade 12 NAEP United States history test (NAEP-USH), we sought to disentangle the interconnections of students' instructional exposure, sociocultural characteristics, and content knowledge within and across schools to better understand the relationships among these pieces and their relative association with students' learning of history.

The Case for History Content Knowledge

Over the last several years, history educators have shifted their prioritization. No longer focused on the simple collection of facts and ideas, they emphasize the "doing of history"—the cognitive practices of investigation, interrogation, and interpretation of the past (Barton, 2012; Lee, 1994; VanSledright, 2011). Emphasis on inquiry privileges the process of history, rather than the narrative. In doing so, inquiry-based pedagogy teaches students to seek evidence to guide their thinking, assess the veracity of source material, and engage in critical thinking. History through inquiry promotes a democratization of ideas essential to a society.

Yet, narrative structures continue to embed the study of history, providing a template for both the contextual and temporal understanding of the past. Furthermore, principles of cognitive theory propose that development of higher-order thinking, such as historical inquiry, is predicated upon (and influenced by) students' frames of reference or schema (Wertsch, 2002). Providing recommendations for the constructivist teacher, Brophy (2006) suggested that in order for students to make authentic connections in their learning, teachers must provide appropriate opportunities to develop knowledge of fundamental content and concepts. Otherwise, teachers run the risk of encouraging ambiguous relativism, leaving their students intellectually adrift. Wineburg (1998, p. 339) argued that it is important for historical interpreters to recognize the context and vocabulary of what they are reading—helping to "constrain" the meaning in order to avoid uninformed inquiry. Evaluating the prominent canon of the United States is consequential for understanding how Americans as a society collectively identify and perceive the past (Reich, 2011; VanSledright, 2011). This collective narrative, while subject to scrutiny (e.g., Howard, 2003), functions as a common frame of reference that allows for meaningful contextualization of source material, the anchoring of inquiry-based instruction, and construction of a common national identity (Reisman, 2012; Reisman & Wineburg, 2008).

In civic education, researchers have made similar claims regarding the importance of a common knowledge base (Galston, 2007; Zhang, Torney-Purta, & Barber, 2012). Specifically, they argue that individuals' knowledge concepts and content regarding the institutions, rights, and processes encourages voting and other forms of active civic participation. Using NAEP civics data, Niemi and Junn (1993, 1998) theorized an interconnected exposure and selection process to explain how and to what extent students retain civic knowledge. They posited that learners' exposure to civics was contingent upon their schooling experience, including exposure to content and pedagogy. Thus, access to curriculum along with optimal delivery of instruction increases students' exposure. Niemi and Junn, however, further contended that exposure to content failed to adequately explain what civic content students retain. Rather, selection of the content mitigates exposure; learners select (i.e. retain) material based upon their interest in the content, motivated by individual sociocultural elements such as race, class, parental involvement, gender, and past experience with the subject.

Influences on Historical Content Knowledge

Similar to civics processes espoused by Niemi and Junn, we posit that students' knowledge attainment in history is an interconnection of schooling experiences and cultural identification. The

following sections highlight the importance of both the sociocultural context and formal instructional experiences as influences on students' history content knowledge.

The Influence of Sociocultural Characteristics

History, as a discipline, attempts to understand the past. Values and cultural norms, which influence how societies understand the past and prioritize historical accounts, shape learners' engagement with content knowledge. VanSledright (2011) and Wertsch (2002) refer to this canonical development, as a freedom-quest narrative, emphasizing a collective national identity of progress. This traditional historical canon helps contextualize and construct historical arguments, while also serving to bind people culturally. Yet, historical interpretation and knowledge are socially constructed (Barton, 2012). How and to what extent individuals engage the past depends, to no small extent, upon their own subjectivity. History education researchers (VanSledright, 2011; Wineburg, Mosborg, Porat, & Duncan, 2007) note that a learner's positionality and sociocultural context within the curriculum influence acceptance of a particular historical narrative. For students to recognize value in the curriculum and remain motivated to learn, discipline-specific research suggests content should be presented as relevant and connected to the lives of learners (Alexander, 2003; Wineburg et al., 2007). Learners whose identities do not reflect the major figures and cultural attributes of history more often fail to connect with the content. Students, unmotivated to engage with curriculum incongruent to their cultural and social norms, retain less formal history knowledge (Chikkatur, 2013; Epstein, 2009; VanSledright, 2008).

Notable scholars of diversity education have critiqued the content of social studies curriculum, holding particular antipathy toward history (Gay, 2003; Howard, 2003; Ladson-Billings, 2003). They argue a lack of race-consciousness in the traditional history canon contributes to skepticism of the discipline among communities of color. Research suggests that Black students' interpretation and acceptance of historical canon varies from their White counterparts. In her study of how Black students interpret history differently from White counterparts, Epstein (2009) found Non-White students take a more critical view of the freedom-quest narrative compared to their White peers, and display outward pride for civil rights struggles and accomplishments. Concomitantly, Chikkatur (2013) found that African American students were more receptive to a curriculum with which they identified. Unfortunately, such episodes within American history canon are more additive than normative, making it difficult for students of color to find their place and identity within the curriculum.

Lack of gender equality in historical narratives is another point of contention. Female students report finding the masculine historical narrative less appealing; thereby, influencing their interest and retention of historical information over years of schooling (Fredrickson, 2004). Research suggests that women are more likely to appreciate history focused on social change, access to democracy, and civil rights (Crocco, 2008). However, akin to the exclusion of Non-Whites, these themes do not resonate in canonical representations of the past, which frequently privilege male-oriented motifs such as war, politics, and power (VanSledright, 2011). Complicated race and gender connections to the past, students from lower socioeconomic backgrounds often lack cultural capital, receive less parental involvement, and maintain lower academic expectations—all of which negatively impact their student achievement (Berliner, 2006; Lee & Bowen, 2006). Future educational goals also serve as predictors of student engagement and interest in content (De La Paz, 2005; Smith & Niemi, 2001).

The Influence of Disciplinary Literacy

History education research suggests students exposed to a variety of instructional modes requiring students to engage text-based resources are more likely to retain content (Reisman, 2012).

Researchers have linked instructional exposure associated with reading, writing, discussion of content, and analysis of documents to “core” teaching practices (Barton & Avery, 2016; Fogo, 2014). These instructional practices are also reflected as essential skills in the Common Core State Standards and manifest as forms of disciplinary literacy (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010).

Disciplinary literacy is the confluence of information, experiences, and abilities possessed by those who create, communicate and use knowledge within a specialized subject area (Shanahan & Shanahan, 2008). Processes to this end include reading, writing, listening, speaking, and thinking critically in ways meaningful within the context of disciplinary work. The aim of disciplinary literacy is to “transform students into disciplinary insiders” who are equipped with the specialized skills necessary to meet the demands and mores of content domains (Shanahan & Shanahan, 2012, p. 11). These specialized literacy skills give students agency when asked to comprehend and derive meaning from complex informational texts prevalent in history.

This meaning-making process requires a text-dependent interplay between reading material, task and reader. More explicitly, disciplinary literacy in history education includes inquiry and author awareness, reading and analysis of various historical source material, writing as an extension of the text engagement, and the use of discussion to elaborate and qualify textual understanding (Giles, Wang, Smith & Johnson, 2013; Moje & Speyer, 2008; Monte-Sano, 2010; Wineburg & Reisman, 2015). It is through these tenants of disciplinary literacy that students engage content in multiple and meaningful ways, connecting content and concepts to working memory. Though few large-scale studies have attempted to ascertain if exposure to such instruction are core practices that optimize students learning potential, modest-scale studies suggest that exposure to these text-dependent methods has the potential to improve students’ understanding and knowledge acquisition in history (Monte-Sano & De La Paz, 2012; De La Paz, et al., 2014; Nokes, 2010; Reisman, 2012; VanSledright, 2011).

Discipline-specific literacy exposure is also associated with test performance. Reich (2009), in his analysis of how students respond to items on the New York Regents Exams, found that students’ discipline-specific vocabulary and test-savvy skills were tied to cognitive reasoning on exam items. He further intimates that exposure to reading and writing within the field potentially improves test acumen on historical content assessments. Similarly, Reisman’s (2012) experimental study of a reading historical documents program found students exposed to targeted reading strategies scored higher on history tests than those who did not receive disciplinary literacy instruction.

It is important to remember that the aforementioned influences on history knowledge attainment, while grouped for the purposes of this study, are interconnected and should not be compartmentalized. Sociocultural characteristics, such as race and class, can have substantial ramifications for the instruction of students; as research notes that students from non-White backgrounds and from lower socioeconomic families are more likely to receive less qualified social studies teachers and substandard instruction (Fitchett, 2010; Pace, 2011; Segall, 2006). Instructional effects are also pronounced among English language learners and students with special needs, many of whom receive substandard history/social studies instruction (Cho & Reich, 2008; Litner & Schweder, 2008; Szpara & Ahmad, 2006). This teaching divide intensifies unequal opportunities to learn for students and discourages authentic access to the curriculum.

Assessing History

Much debate has centered over the purposes and practices of historical assessment. Seixas and Ercikan (2015) noted there are more multiple ways to appraise historical understanding

including knowledge retention, disciplinary procedures, and perspective-taking. They have contended that assessment multidimensionality strengthens the discipline as long as tests accurately measure what they claim to assess. Conflict regarding assessment has arisen in acknowledging what history educators should privilege in their assessments, and what students ought to be able to know and do. Because history education in the United States has traditionally focused on a canonical transmission of the past, assessments have primarily focused on memorization and factual recall (VanSledright, 2014). These exclusively multiple-choice tests often forsake construct validity (measuring what they purportedly assess) for responder reliability and grading efficiency. Moreover, they are often vulnerable to test-taking gamesmanship by students (Reich, 2009). The growing emphasis of disciplinary thinking in history education has led to challenges and criticisms of these traditional assessments. Advocates of a more skill-based assessment of history (Smith & Breakstone, 2015) have challenged conventional history tests, suggesting that they fail to gauge students' disciplinary understanding (i.e., the "doing" of history).

Further complicating history assessment policies are the lack of federal mandates for testing history or other social studies content (Fitchett et al, 2014a). Thus, it is difficult to evaluate students' understanding of history across state lines and consequentially compare curricular policy and practices in history teaching and learning. The National Assessment of Educational Progress U.S. history test (NAEP-USH), described in more detail in subsequent sections, provides a viable alternative for examining history learning and teaching in the United States. In assessing both students' content knowledge and analytical skills (Lazer, 2015), NAEP-USH offers a national snapshot of history learning in the nation. Furthermore, it can be used to compare sub-groups of students to determine how sociocultural influences and exposure to various instructional practices (e.g., disciplinary literacy) are associated with student learning.

School Policy Context and History Content Knowledge

Within the study of history and the social studies, research also acknowledges that the aforementioned sociocultural and instructional exposure influences are consequential at the school-level (Epstein, 2009; Levinson, 2012; Pace, 2011; Segall, 2006). School climate and the surrounding community culture influences how students prioritize and engage historical content knowledge. Moreover, testing culture and accountability requirements constrain history teacher decision-making (Sayre et al., 2013) and can promote rote, lower-level instruction (Volger, 2006). Social educators have lamented that accountability and high-stakes testing have hijacked curriculum, converting learning contexts into cram schools focused on test preparation and offering minimal learning engagement and even fewer opportunities for complex, higher level thinking (Volger & Virtue, 2008). However, competing studies suggest, that while accountability policies have shifted *what* history students learn, little evidence has indicated that it actually changed *how* they learn (Grant, 2001; Patterson, Horner, Chandler, and Dahlgren, 2013; van Hover, 2006). Hence, others argued, accountability and testing hold little influence over history teaching.

Given this evidence, one might conclude that there has been little return on investment for educational policy promoting testing and curriculum standardization in history education; however, previous research from Fitchett et al. (2014a, 2014b) has hinted otherwise. Analysis of elementary social studies indicated that testing was associated with increased time spent on the subject. Moreover, testing, while associated with lower perceptions of autonomy among practitioners, was a more consistent predictor of instructional time than teachers' perceived classroom control, perhaps suggesting that accountability does increase access to time spent on history/social studies course content. Few studies in history/social studies have taken the next step—examining the association between state-level testing and overall student achievement. Therefore, it is necessary to delve into

the broader (and more contentious) literature on relationships between accountability mandates and student achievement.

Among wider educational research and policy circles, there is considerable debate over the efficacy of accountability and state testing as appropriate measures for improving student achievement. Amrein and Berliner (2002), in their archival time-series analysis of NAEP math and reading scores, found little evidence that high stakes testing policies at the state level were associated with higher achievement. In contrast, Rosenshine (2003), examining the same NAEP data with a control group design, concluded that states with testing policies correlated with higher scores on average compared to states without high-stakes testing. He further posited that these gains were not due simply to test preparation, but rather a concerted effort to referee the curriculum content taught, which often aligns with NAEP standards. In a rebuttal, Amrein-Beardsley and Berliner (2003) incorporated Rosenshine's control design and found that, while states with high-stakes tests performed better than non-testing states on fourth grade math NAEP, results in reading and other grade bands were inconclusive. Moreover, they concluded that NAEP sampling frames often exclude students with learning disabilities and exceptionalities, thus making state comparison unreliable.¹ While important findings, these competing studies focused at the state-level, without sufficiently accounting for variance at the building- or student-level. In response to these earlier studies, Dorn (2006) argued the need for more advanced, multilevel analysis of NAEP data to better examine teaching and learning across state-policy contexts.

Other robust educational policy analyses find increased accountability policies are associated with higher NAEP scores (Braun, 2004; Carnoy & Loeb, 2002; Hanushek & Raymond, 2006), implying higher standards and accountability measures such as testing externally motivate students and teachers, while regulating instructional content. From a cognitive perspective, research suggest testing and constant evaluations serve as potential tools for reinforcing knowledge, skills, and concepts. Carpenter and colleagues (2009) contended that testing has potentially positive effects on students' long-term historical memory. Sousa (2011) affirmed from neurological science, that testing is an effective measure, when used to promote self-regulation, for gauging sustained learning outcomes. However, from a policy perspective, research has not examined the potential influence of testing and accountability policy on what students learn in history and the social studies.

Purpose of the Study

Whereas the majority of the previous research on historical understanding and achievement has focused on either sociocultural or instructional exposure influences, few studies have examined their complex interrelatedness on a large scale while also accounting for complex between-school characteristics including the state policy climate. In this study, we analyzed the relationship among various sociocultural and instructional exposure influences on US content knowledge outcomes for 12th graders using data from the National Assessment for Educational Progress US History Assessment (NAEP-USH).

Four research questions guided our exploratory analyses:

- 1) To what extent are student-level sociocultural characteristics associated with US history content knowledge as measured by NAEP-USH?

¹ NCES has taken steps to insure greater inclusion on NAEP in subsequent years. Findings from these policies can be found here: National Center for Educational Statistics (2008). *Measuring the Status and Change of NAEP Inclusion Rates for Students with Disabilities: Research and Development Report* (NCES 2009-453). Washington, DC: US Department of Education.

- 2) To what extent are student-level instructional exposure variables associated with US history content knowledge as measured by NAEP-US?
- 3) To what extent are school-level demographics characteristics and accountability contexts associated with US history content knowledge as measured by NAEP-US?
- 4) To what extent do school-level characteristics and contexts serve as moderating effects on the racial achievement gap associated with US content knowledge as measured by NAEP-US?

Method

Participants

We analyzed NAEP-US scores of 12th grade, public school students ($n_{student}=10,890$) nested in schools ($n_{school}=410$).² Small amounts of missing data were present for some of the student demographic characteristics. Multiple imputation was used to replace the missing data. Public schools were specifically chosen because of the inherent uniformity of curriculum and accountability policies found among public schools within a state compared with variable conditions placed upon teachers and students in private schools. Grade 12 was examined exclusively because it was posited that historical knowledge acquisition occurs throughout schooling experience; accordingly, the 12th grade assessment represents an accumulation of knowledge.

Materials

NAEP US history assessment. For the dependent variable, we used student achievement estimates from a nationally representative sample of students who took the 12th grade 2010 National Assessment for Educational Progress US History Assessment (NAEP-US) to assess US history content knowledge (National Center for Education Statistics, 2011). Established in 1994, NAEP-US is used primarily for research purposes by the U.S. Department of Education and outside researchers to examine historical knowledge across student sub-populations (Lazer, 2015). The federal government administers the test approximately every four years to approximately 10,000 students at three different grade levels, grades 4, 8, and 12. NAEP-US is a low-stakes assessment. The results from the test are not used for student grading purposes or to evaluate teacher performance. The value of the assessment lies in its national sampling frame and the generalizability of the results. Therefore, NAEP-US serves as the most comprehensive cross-sectional assessment of US history content and conceptual knowledge in the nation.

Test items are constructed evenly around a series for four themes: change and continuity democracy; the gathering and interactions of peoples, cultures and ideas; economic and technological challenges; and the changing role of American in the world (NAGB, 2010). Eight historical period are representative of the test including: pre-colonial American history, American Revolution, the U.S. Civil War, emergence of modern American, World War II, and contemporary issues. Item development and content validation processes included groups of history teachers, history professors, and history education professors. Their recommendations were presented to the NAEP Assessment Governing Board, who made final decisions on item inclusion and content

² Students were included where the number of students within a school building ≥ 5 . All samples sizes were rounded to the nearest 10 in compliance with NCES data reporting procedures.

(NAGB, 2010). NAEP-USC at grade 12 includes 110 multiple choice, 36 short open-ended, and 13 extended open-ended items (Lazer, 2015).³

NAEP test developers claim the assessment measures cognitive processes of “historical knowledge and perspective” and “historical analysis and interpretation” (Lazer, 2015; p. 147). Given this claim of breadth and depth of historical thinking, test developers encounter two competing dilemmas: how to measure breadth of content knowledge, while also measuring analytical skills. However, unlike traditional standardized assessments, NAEP-USC uses matrix sampling, whereby the test is divided into 10 blocks. Students complete two 25-minute blocks of the aggregate exam for a total test time of 50 minutes. Each block is paired with all other blocks at least once. These content-themed blocks (i.e. Great Depression and World War II) allow NAEP-USC to measure both students’ historical knowledge and, to an extent, their ability to use discipline-specific cognitive processes across content domains. Students only take a portion of the entire assessment (two of ten blocks) and avoid potential test fatigue associated with administering a long, burdensome survey of content knowledge to students. NAEP-USC also attempts to counter issues of reliability associated with offering a narrow set of interpretative questions, requiring first-order historical knowledge that a student might (or might not) already possess.

Because a student only takes a portion of the overall assessment, they do not receive individual scores. Rather, individual outcomes are turned into predictive scores based on student group characteristics. Predictive scores are statistically calculated into plausible values (PVs). These values are created by a multi-stage process. Item-response theory modeling is used to generate the mean and variance of the expected distribution of scores for subgroups of students. Researchers are provided five PVs for each student which represent randomly sampled values from the expected distribution of the subgroup to which the student belongs. These PVs are standardized into composite scores with a range of 0 to 500.

The results of NAEP-USC and their usefulness to history educators has been widely scrutinized and debated. Detractors suggest that NAEP scores do not capture the complexity of historical thinking (Breakstone, 2014; Rothstein, 2004); however, others argue that tests like NAEP provide a baseline of students’ understanding of the common historical narrative, a narrative up for critique but essential for contextualization (Reich, 2009, 2011; VanSledright, 2011; Wineburg, 1998). Given the criticism leveled against using NAEP-USC as a representative gauge of higher-order historical thinking, interpretation of student outcomes was tempered. This study defined NAEP-USC results as a measure students’ content knowledge acquisition and first-order reasoning rather than a range of more complex historical understanding processes espoused by previous researchers (cf. Barton, 2012; VanSledright, 2011).⁴

Sociocultural characteristic predictors. In addition to the assessment items, NAEP-USC also provides data on student demographics and instructional exposure. Informed by early research sociocultural indicators served as independent variables in this study. They included dummy-codes for gender and graduation, free/reduced lunch status. Previous analyses of NAEP-USC found that

³ Sample items from the NAEP-USC 2010 can be found at: http://nces.ed.gov/nationsreportcard/pdf/demo_booklet/10SQ_Gr12_CHG.pdf

⁴ One common criticism of NAEP-USC has been because the lack of long, essay-type questions similar to those posed on Advanced Placement exams for measuring students’ historical interpretation skills. NAEP-USC avoided using these items because students might not have the necessary pre-requisite reading and writing skills necessary to complete the item; thereby increasing the amount of missing achievement data (Lazer, 2015). Such items also take substantial time to complete and grade. The short and extended open-ended items were viewed as an imperfect compromise. They provide students opportunities to offer authentic interpretation of text, while also minimizing item response time to under 10 minutes.

students who identified as Black, Hispanic, or Native American followed similar trends on NAEP-USH test performance (Heafner & Fitchett, 2015). Therefore, race identification was dichotomized into a non-White, non-Asian indicator. Because cultural capital and educational access are multifaceted, we included parental college enrollment, plans to attend a four-year college, and need for special accommodations (designated Individual Education Plan and/or 504 plan) as sociocultural control variables.

To control for biased estimates from the low-stakes nature of NAEP-USH, a variable for student interest in US History (a standardized Likert-type scale) was incorporated into the models (Lazer, 2015; Niemi & Junn, 1998) as a sociocultural control variable. In addition, students who have not completed US history are potentially disadvantaged on the assessment; hence, they experience a limited opportunity to learn the content. Students' taking US history in 12th grade was included as a covariate. Moreover, controlling for measures of prior academic achievement reduces the likelihood of reverse causality; whereby, students' achievement determines classroom practice (Podgursky, 2002). NAEP-USH, much like other cross-sectional NAEP datasets, does not include direct measures of prior achievement (Dorn, 2006).⁵ In lieu of prior academic performance, we also used students' age as a covariate among the sociocultural indicators, positing that students older than the modal age were either grade retained or were slower to mature academically (for other examples see Schmidt, Burrough, Zoido, & Houang, 2015).

Instructional exposure predictors. The grade 12 NAEP-USH provides student-reported data on their instructional exposure, which were incorporated as independent variables in this study. Because teaching and learning in history is not specific to a singular instruction type (Fallace, 2010; Levstik, 2008), we conducted a principal axis factor analysis with orthogonal rotation to determine whether instructional exposure items loaded onto similar constructs. Results indicated that scales loaded onto two factors. One factor, labeled Text-Dependent Instruction, included items measuring frequency of discussion of materials, readings from textbook and other sources, engage historical documents, and write short answers. The second factor, labeled Multimodal Instruction, included items measuring frequency of work on group projects, class presentations, writing reports, going on field trips, listening to information online, and use books or computers in library for schoolwork. One item, frequency of testing and quizzes, was not included in the factors due to lack of fit.⁶ Both factors were found to exhibit adequate internally consistent reliability within the sample ($\alpha_{text-dependent}=.743$; $\alpha_{multi-modal}=.727$). The unidimensional factor loading patterns were similar to previous analysis of NAEP-USH data (Heafner & Fitchett, 2015; Smith and Niemi, 2003). Conceptually, the factors aligned with existing literature surrounding students' learning of history. As noted above, the teaching and learning of history point toward the importance of disciplinary literacy, which includes the reading, writing and discussion of historical texts. Furthermore, given the potential to motivate students, efforts to "bring history alive" through varied experiential activities are popular in today's classrooms (cf. Wright-Maley, 2015). Therefore, it was determined that the two-factor solution adequately reflected (conceptually and statistically) students' instructional exposure to history. The two factors scaled as composite scores, standardized, and included as independent variables. In the classroom, the exposure to instruction is rarely dichotomized into either Text-Dependent or

⁵ NAEP-USH includes only cross-sectional data. Prior knowledge on the NAEP has previously been examined using High School Transcript Study data provided by NCES. Smith and Niemi (2001) used this data in an earlier examination of NAEP-USH. However, those data were not available for the 2010 NAEP-USH.

⁶ Previous analysis (Heafner & Fitchett, 2015) indicates that this item was not significantly associated with NAEP-USH scores. For the sake of model parsimony, we did not include it in this study.

Multimodal instruction. History teachers are likely to use a variety of instructional strategies (Fallace, 2010; Levstik, 2008). An interaction term was created to examine how students' exposure to Text-Dependent instruction as moderated by Multimodal instruction was associated with NAEP-US. A scale score was created from the interaction term and standardized.

Within instructional exposure variables, Advanced Placement US History class enrollment (APUSH) was operationalized as a substitution for prior student achievement and as a control for instructional exposure predictor. Entry into AP coursework is typically contingent upon a successful academic record and previous research also notes that students taking APUSH courses experience instruction differently than students in non-AP learning environments (Sayre et al., 2013; Smith & Niemi, 2001). We also included an indicator of online instruction as a control variable, positing that online teaching and learning environment are demonstrably different than traditional classroom environments. Descriptive statistics of the sample are included in Table 1.

Table 1

Descriptive Statistics of NAEP-US Participants ($n_{students}=10,890$ in $n_{schools}=410$)

Level	Dependent Variable	Mean (SD)
	Mean PV score for US History	285.43 (32.28)
Independent Variable		
Student Level	Female	50%
	Non-White, Non-Asian Minority	40%
	Free-Reduced Lunch	37%
School Level	>75% School is F/R	14%
	% School Black	17%
	% School Hispanic	19%
	Test Social Studies at all Grade Bands	27%
	Test Social Studies at Middle and High School	30%
	Test Social Studies at the High School Only	7%
	No Social Studies Test	36%

School-level demographic characteristics and policy context. Because NAEP-US samples regionally rather than statewide, it was not possible to analyze a three-level model to explore state to school to student effects. Rather, researchers incorporated state-level accountability data as building-level contextual variables; positing state mandates would have similar impact within schools of the same state. To measure US History assessment and testing policy, Education Week's 2010 *Quality Counts* survey of state-level testing policy was employed (Executive Summary, 2010). From available data, three state-level indicators assigned to the school-level were created: test social studies at high school only, test social studies at middle and high school, and test social studies at all three

grade levels (elementary, middle and high school). Previous studies have used *Quality Counts* to designate similar states' policy contexts (Fitchett et al., 2014a, 2014b). NAEP-USH provided quartile measures of school percentage free-reduced lunch and school percentage planning to attend a college/university. In this study, researchers included indicators of greater than 75% free-reduced lunch and less than 25% of students college-bound to examine the effects of concentration poverty and lower than average post-secondary academic readiness on average history knowledge achievement. In addition, school level percentages of Black and Hispanic enrollment were included as school-level predictors. School-level indicators of charter school identification and urbanicity were included as building-level control variables.⁷ Table 2 includes a list of all predictor and control variables modeled in this study.

Table 2
Predictor and control variable list

Variable	Description
<u>Sociocultural predictors</u>	<u>Student-level</u> Female (comparison category: Male) Non-White, Non-Asian Free/reduced lunch eligible
<u>Sociocultural controls</u>	Parents did not attend a 4-year college Do not plan to attend a 4-year college Interest in US history Currently taking US history English language learner indicator Learning accommodation (IEP and/or 504) Older than modal age
<u>Instructional exposure predictors</u>	Text-dependent instruction Multimodal instruction Text-dependent instruction x multimodal instruction
<u>Instructional exposure controls</u>	Online instruction Enrolled in/Taken APUSH
<u>School characteristics predictors</u>	<u>School-level</u> <%25 of student attend a 4-year college >75% eligible for free/reduced lunch % School Black % School Hispanic
<u>Accountability contextual predictors</u>	Text social studies at: All grades, Middle and high school only; High school only (comparison category: No test)
<u>School-level controls</u>	City (comparison categories: Rural, Suburban) Charter (comparison category: Not a charter)

Note: All variables were included in the HLM model. Coefficients for predictors were included Table 2. Full tables, which include results of control variables, are available upon request from the authors.

⁷ NAEP school context variables only provide continuous variables of % Black and % Hispanic at the school. All other items were categorical. Dummy variables were created.

Analytical Procedure

We conducted multilevel modeling (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2004) using HLM software to document relationships among student sociocultural characteristics, instructional exposure, and school-level context and US history content knowledge. The plausible values algorithm in HLM was used to model the dependent variable for each student. Models accounted for the nesting of teachers in schools and determined if there were any significant building-level effects associated with NAEP-USC outcomes. Five models guided the analyses. To answer research question one, Model 1 examined the association between student-level sociocultural variables and controls on NAEP-USC outcomes. To answer research question two, Model 2 introduced student-level, instructional exposure variables and controls into the model to determine their unique contribution to the variance in history content knowledge. Model 3 included the interaction term and furthered our exploration of research question two. To answer research question three, Model 4 included school-level contextual and state policy variables to determine their unique contribution to between building variance of historical content knowledge. To answer research question 4, model 5 examined the extent that school building characteristics moderated the race effect (non-White, non-Asian) on NAEP-USC. The composite scores and interaction terms were group-mean centered at Level 1. The percentages of Black and Hispanic students were grand mean-centered. Student- and school-level sampling weights were used to provide more accurate coefficients. Robust standard errors were used to account for the complex sampling design.⁸

Results

The Intraclass Correlation Coefficient (ICC) for the unconditional model was .200, indicating that 20.0% of the variance in history knowledge outcomes was found between schools and 80.0% between students within schools; justifying the use of multilevel modeling. Table 3 presents results from each of the three analyses associated with the student level effects for the full sample. Table 4 presents the results of the school level analyses. Model 1 sociocultural predictors and covariates accounted for approximately 30% of the within-school variance associated with NAEP-USC (historical content knowledge). The inclusion of instructional exposure-type predictors and covariates in Models 2 through 5 accounted for an additional 9% of the within-school variance associated with historical knowledge acquisition; suggesting that approximately 39% of the variance within-schools could be explained by the analysis.

⁸ The HLM equations for the full model:

Level-1 Model

$$\begin{aligned} HRPCM_{1j} = \beta_{0j} + \beta_{1j}*(FEMALE_{ij}) + \beta_{2j}*(NONWHTAS_{ij}) + \beta_{3j}*(NOPAR_CL_{ij}) + \beta_{4j}*(ELL_{ij}) + \\ \beta_{5j}*(FRLUNCH_{ij}) + \beta_{6j}*(NOKID_CL_{ij}) + \beta_{7j}*(HISTNOW_{ij}) + \beta_{8j}*(INTEREST_{ij}) + \beta_{9j}*(OLDER_{ij}) + \\ \beta_{10j}*(ACCOM1_{ij}) + \beta_{11j}*(APUSH_{ij}) + \beta_{12j}*(ONLINE_{ij}) + \beta_{13j}*(TD_{ij}) + \beta_{14j}*(MM_{ij}) + \beta_{15j}*(TDxMM_{ij}) + r_{ij} \end{aligned}$$

Level-2 Model

$$\begin{aligned} \beta_{0j} = \gamma_{00} + \gamma_{01}*(COLL_25_j) + \gamma_{02}*(FR_75_j) + \gamma_{03}*(CITY_j) + \gamma_{04}*(CHARTERY_j) + \gamma_{05}*(SSCHBLK_j) + \\ \gamma_{06}*(SSCHHSP_j) + \gamma_{07}*(ALL_TEST_j) + \gamma_{08}*(MDHS_TES_j) + \gamma_{09}*(HS_ONLY_j) + u_{0j} \\ \beta_{1j} = \gamma_{10} \\ \beta_{2j} = \gamma_{20} + \gamma_{21}*(COLL_25_j) + \gamma_{22}*(FR_75_j) + \gamma_{23}*(CITY_j) + \gamma_{24}*(CHARTERY_j) + \\ \gamma_{25}*(SSCHBLK_j) + \gamma_{26}*(SSCHHSP_j) + \gamma_{27}*(ALL_TEST_j) + \gamma_{28}*(MDHS_TES_j) + \gamma_{29}*(HS_ONLY_j) + u_{2j} \end{aligned}$$

At Model 1, approximately 67% of the between school variance in NAEP-USC performance was explained. In Models 4 and 5, an additional 9% of the variance was explained from the school characteristics. Overall, the analyses accounted for approximately 76% of the between school variance in performance.

Research Question 1: Influence of Sociocultural Indicators on US History Content Knowledge

Findings indicate that a proportion of the variance in students' knowledge of historical content as measured by NAEP-USC can be attributed to sociocultural variables (see Table 3). Results confirm previous smaller-scale analyses that race, class, and gender are associated with variability in students' history content knowledge. When controlling for other variables, Non-White/Non-Asian minority associated with approximately a one-third a standard deviation lower on NAEP-USC compared to their counterparts. Findings also reveal a gender gap, whereby, male students consistently out-perform female students on NAEP-USC by approximately .20 of a standard deviation. Students eligible for free-reduced lunch and non-college bound students are also associated with lower performance in US history content knowledge.

Research Question 2: Influence of Instructional Exposure on US History Content Knowledge

While the inclusion of exposure-type variables in the analyses are a modest contribution to the overall model estimate, the analyses reveal important findings for social studies and history educators (see Table 3). In Model 2, each standard deviation increase in text-dependent instruction (including reading, writing, and discussing historical content) is associated with approximately 6.5 points of higher performance on NAEP-USC (or almost a quarter of a standard deviation in achievement). Conversely, each standard deviation increase of multimodal instruction (e.g. group projects and field trips) is associated with the inverse, which equates to a 7-point drop in performance on NAEP-USC. Since instruction in history/social studies is rarely bifurcated between these typologies, we included an interaction term in Model 3. Results suggest that multimodal instruction inversely moderates the positive effects of text-dependent instruction. The interaction of text-dependent and multimodal instruction was associated with slightly more than one quarter of a standard deviation decrease in student performance on the NAEP. Thus, for each standard deviation increase in interaction (more exposure to both) the higher performance associated with text-dependent instruction is offset by multimodal instruction.

Research Question 3: Influence of School Characteristics and Accountability Policy on US History Content Knowledge

Models 4 and 5 accounted for a substantial amount of between school variance (approximately 76%). Indicators of school-wide academic achievement (<25% of the school students attending a 4-year college), the percentage of Black students, and state testing policy were the statistically significant school-level variables. Every 10% increase in the number of Black students in schools was associated with 1.8 point average lower test performance. Testing policy, employed as a proxy for accountability structures, was associated with higher average building-level performance on NAEP-USC. Students who were tested at the high school level were associated with a 3 to 4 point average increase on the test. Moreover, testing across all grade bands was associated with significant and approaching significant increases when controlling for other building level characteristics.

Table 3

Level I Fixed Effects Estimates for Models of Selection-Exposure Predictors of US History Content Knowledge (n=10,890)

Parameter	<u>Model 1</u>	<u>Model 2</u>	<u>Model 3</u>	<u>Model 4</u>	<u>Model 5</u>
	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
Intercept	307.80 (1.13)***	306.60 (1.13)***	303.58 (1.13)***	302.42 (1.37)***	303.54 (1.48)***
<i>Sociocultural indicators</i>	Female	-6.08 (0.77)***	-6.96 (0.71)***	-7.00 (0.71)***	-6.97 (0.70)***
	Non-White, Non-Asian	-13.66 (0.88)***	-12.16 (0.85)***	-12.15 (0.85)***	-10.38 (0.86)***
	Free/Reduced Lunch	-4.76 (0.81)***	-4.06 (0.79)***	-4.10 (0.79)***	-3.68 (0.78)***
<i>Instructional Exposure indicators</i>	Text-Dependent instruction (TD)		6.58 (0.46)***	13.62 (2.12)***	13.61 (2.12)***
	Multimodal instruction (MM)		-6.95 (0.46)***	0.13 (2.09)	0.02 (2.08)
	Interaction (TD x MM)			-9.16 (2.66)***	-9.13 (2.65)***

Note. Coefficients displayed for independent variables of interest. HLM models included full school and teacher control variables.

***- $p < .001$, **- $p < .001$, *- $p < .05$, †- $p < .10$

Table 4

Level II Random Effects Estimates for Models of Selection-Exposure Predictors of US History Content Knowledge (n=410)

Parameter	<u>Model 4</u>	<u>Model 5</u>
	γ (SE)	γ (SE)
β_0 NAEP-USN		
<25% of Students Attending 4-year College	-5.69 (1.39)***	-6.50 (1.58)***
>75% School is F/R	-2.20 (1.92)	-0.32 (3.20)
% School Black	-0.18 (0.03)***	-0.07 (0.04) †
% School Hispanic	-0.03 (0.03)	-0.06 (0.04)
Test Social Studies at all Grade	3.06 (1.50)*	2.90 (1.72) †
Test Social Studies at Middle and High School	1.93 (1.61)	1.87 (1.77)
Test Social Studies at the High School Only	3.06 (1.48)*	4.35 (1.64)**
β_{2j} Non-White, Non-Asian		
<25% of Students Attending 4-year		1.87 (1.71)
>75% School is F/R		-1.79 (3.32)
% School Black		-0.16 (0.04)***
% School Hispanic		0.01 (0.05)
Test Social Studies at all Grade Bands		-0.95 (2.18)
Test Social Studies at Middle and High School		0.13 (1.86)
Test Social Studies at the High School Only		-3.15 (3.32)

Note. Coefficients displayed for independent variables of interest. HLM models included full school and teacher control variables.

***- $p < .001$, **- $p < .001$, *- $p < .05$, †- $p < .10$

Research Question 4: Modeling the Achievement Gap

As a final step (Table 4, model 5), we measured the persistent racial achievement gap that was found across models 1 through 4. Specifically, we examined the extent that building-level characteristics moderate the gap between Non-White, Non-Asian students and their classmates. Results from Table 4 indicated that each 10% increase in the percentage of Black students within the school was associated with a 1.6 point lower building-level average among Non-White, Non-Asian students on the NAEP-USH grade 12 test performance as compared to their counterparts. To further illustrate, these findings researchers graphed results from the study. Figure 1 indicates that among Non-White, Non-Asian (NONWHTAS) students and their classmates, text-dependent instruction was associated with increased test performance.

Figure 2 illustrates that NONWHTAS with high frequency text-dependent instruction outperformed White, Asian students with lower levels of text-dependent exposure when nested in schools with low to moderate Black student enrollment. However, schools identified as predominately Black were associated with lower average NAEP-USH performance for both race categories, even when accounting for text-dependent instruction. Lower performance was most pronounced among NONWHTAS students. Findings suggested that a systemic underperformance on NAEP-USH within predominately Black schools affected students across racial classification.

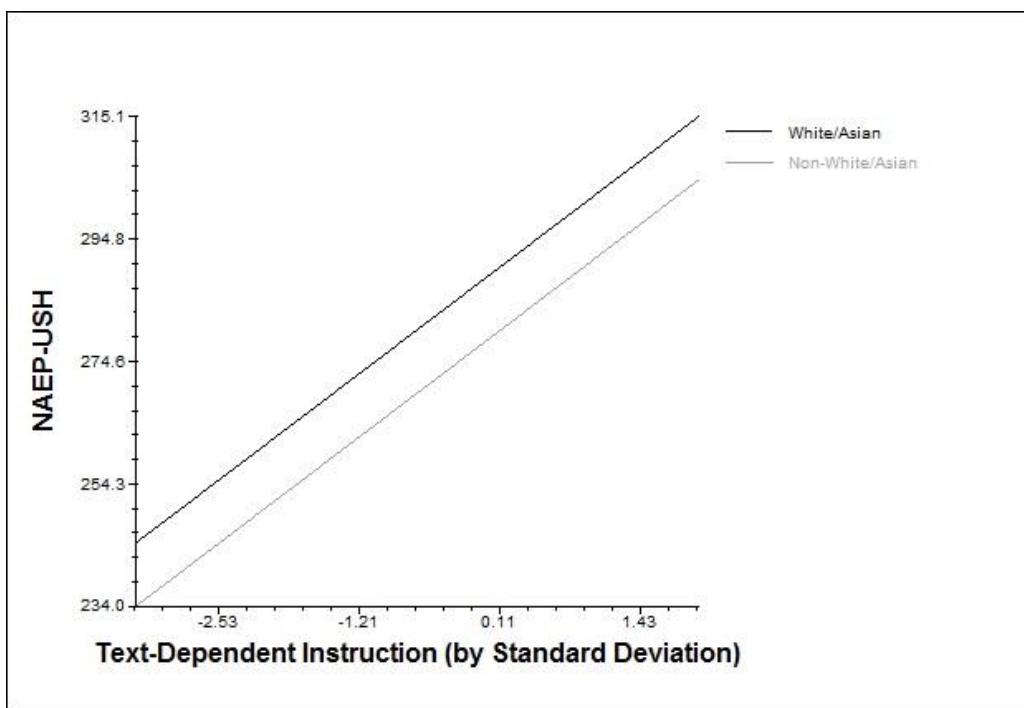


Figure 1. Association between text-Dependent Instruction across NAEP-USH Performance and Race

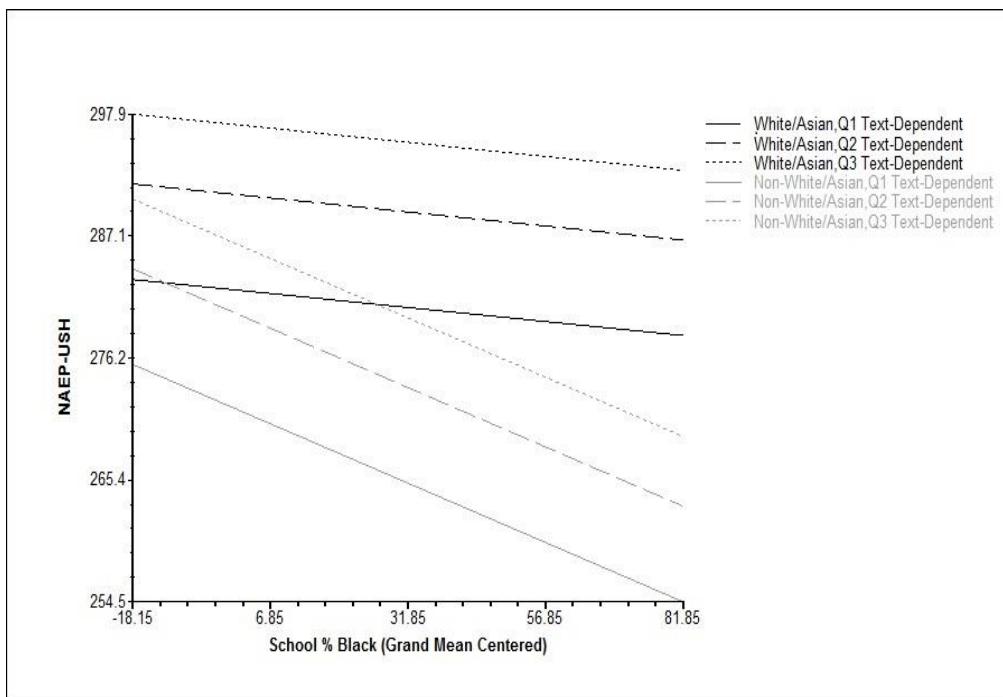


Figure 2. Student Achievement Gap on NAEP-USH as a Function of Text-Dependent Instruction and Concentration of Black Students within the School

Discussion and Implications for Policy and Practice

Providing a large-scale examination of students' history content knowledge from both sociocultural and instructional exposure dimensions suggests that while race, class and gender, serve as dominate predictors of student achievement on history assessments, exposure to text-based instruction makes a positive contribution to what and how students learn. Moreover, building-level demographics and policy context play a significant role in students' between-building achievement level and potentially moderate students' within-building historical content knowledge. These results offer the following implications for policy and practice of history education: a) the importance of acknowledging and countermanding the history knowledge achievement gap, b) the potential for text-dependent instruction as a high leverage instructional practice toward disciplinary literacy, and c) the role of accountability in historical content knowledge.

Acknowledging the History Content Knowledge Achievement Gap

Findings affirm a history content knowledge achievement gap related to the schooling environment and call into question how student communities with non-White, less affluent backgrounds choose to integrate the historical narrative content representative of NAEP-USH. Previous studies offer plausible explanations. Research suggests that the official canon, with an emphasis on White males, is not representative of the diverse population of U.S. schools (Brown, 2011; Cornbleth & Waugh, 1995). Lacking historical positionality beyond the marginalized or the victim, non-White students are more likely to look upon such a narrative with skepticism and dissonance (Chikkatur, 2013; Epstein, 2009; VanSledright, 1998). Consequently, they find learning history/social studies boring, inversely affecting their motivation to learn historical material (Stodolsky, Salk, & Glaessner, 1991; Tanaka & Murayama, 2014). Importantly, race and class should not be conflated. Results indicate, when controlling for race, students who were eligible for free or

reduced lunch were associated with lower NAEP-USC. Access to the official curriculum becomes a fundamental consideration in defining students' opportunities to learn. Given this study's findings, perhaps the official canon perpetuated in standards, curricula, and assessments across U.S. schools deserves further scrutiny. When accounting for other aspects of a students' cultural capital (including socioeconomic status and academic track), non-White, non-Asian students still significantly underperformed, suggesting a dissonance between how students of color and White students perceive history. Perhaps greater emphasis needs to be placed on supporting U.S. history curricula that are inclusive rather than additive and pluralistic rather than monochromatic.

Findings also indicate a discrepancy in historical content knowledge between male and female students. The gender gap remained persistent and aligned with a longstanding gender difference in US History (Zwick & Erckian, 1989), in which males often prefer social studies content (including history) compared to females. This lack of motivation and efficacy toward the study of history is associated with a curriculum that prioritizes male-dominated spheres of politics and war (Crocco, 2006, 2008). If the NAEP does represent the nationalized narrative of the past, then perhaps the gender effect suggests a persistent and over-arching marginalization in how women and feminine topics are included within the history curriculum (Monaghan, 2014; Schmeichel, 2015). Findings from the study imply that additional work at the teaching and curricular level is necessary to improve gender equity and representation in the US history curricula.

Building-level analyses, confirm that the racial achievement divide is not only within, but also between schools—a finding uncomfortably comparable to Levinson's (2012) study of civic understanding. The school percentage of Black students was inversely proportional to average test performance among all students. Modeling the achievement gap within race, results from the analyses found the gap widens when the concentration of Black students is higher. These gaps might be partially explained by the lack of resources and adequate school staffing afforded poor and non-White student populations. Research conducted using the NCES Schools and Staffing survey also suggests that students in predominately Black-enrolled schools are more likely to have a teacher who lacks a teaching license or certification in a social studies/history-related area (Fitchett, 2010). Moreover, the turnover rate, lack of adequate staffing, and dearth of resources in high minority schools is well-documented (Borman & Dowling, 2008; Boyd, Lankford, Loeb, & Wyckoff, 2005). A climate not supportive of the varying needs of history learners can have negative consequences for educational achievement (Ronfeldt, Loeb, & Wyckoff, 2013; Salinas, 2006). The reemergence of segregated schools coupled with a growing socioeconomic gap also contributes to school environments that lack teaching expertise and sufficient resources (Clotfelter, Ladd, & Vigdor, 2006; Mickelson, 2001; Rumberger & Parlardy, 2005). Such schools struggle to meet academic performance expectations of schools with greater fiscal resources. Further research is necessary to explore the between-school differences in how students' understand history and what content knowledge they retain. Furthermore, additional study is needed to counter what Gutiérrez (2008, p. 357) refers to as a "gap gazing fetish," whereby researchers and academics fixate on large achievement discrepancies between White and Non-White students. Avoiding deficit thinking necessitates research that targets outlier exemplars, students of color who succeed in history classrooms. By understanding the characteristics of success, rather than highlighting deficiency, ambitious history education can promote practices that bolster underperforming student populations.

Encouraging Policy for the Practice of Text-Dependent Instruction

Though the majority of the variance within schools was attributed to sociocultural qualities, findings from this study indicate that instruction matters. The use of text-dependent instructional strategies, which includes using various source materials, is associated with greater acquisition of

content knowledge—providing large-scale support of existing historical epistemologies (Guthrie, Klauda, & Ho, 2012; VanSledright, 2011; Wineburg & Reisman, 2015). Given these findings, perhaps it is time for history education to prioritize disciplinary literacy as a high leverage practice. As defined by teacher education researchers (e.g., Ball & Forzani, 2011; Lampert, 2009), high leverage practices are instructional strategies that teachers employ for substantial learning gains. From a cognitive perspective, domain learning models champion discipline-specific reading, writing, and discussion of content to support students' understanding of the past (Maggioni, VanSledright, & Alexander, 2009; VanSledright, 2012). Furthermore, findings from the current study bolster smaller-scale qualitative and experimental results suggesting that exposure to text-dependent instruction (an emphasis of domain learning in history) has positive implications for how students learn historical content (De La Paz, 2005; Monte-Sano, 2012; Reisman, 2012). While previous research has examined teachers' perspectives on core history teaching practices (cf. Fogo, 2014), the positioning of high leverage practices has rarely been explicitly supported in broader history/social studies literature and policy.

The C3 Framework can potentially help support practices that promote discipline-specific literacy skills in history and other social sciences. Drawing from the Common Core Anchor Standards (2010), the framework provides a useful inquiry arc, which privileges disciplinary literacies within historical context. Reading complex texts and engaging in critical discourse through narrative formation enables students to articulate learning in sophisticated ways that mirror the high-level learning demands and career expectations of contemporary society. In an area of faddish educational reforms, teachers and teacher educators should provide students (and preservice teachers) opportunities to engage in meaningful, content-rich activities that support reading and writing in the field. These potentially high leverage instructional practices can positively promote historical knowledge (Moje, & Speyer, 2008; Monte-Sano, 2010; Reisman, 2012) and address student learning outcome differences (De La Paz, 2005; Swanson, et al., 2015). The Stanford History Education Group's (n. d.) *Reading Like a Historian* program is another promising educational tool for engaging students in discipline-specific reading and writing. The program advocates many of the same instructional strategies associated with text-dependent instructional exposure.

Conversely, findings suggest multimodal instruction over-exposure is associated with a decrease in historical content knowledge. While engaging modes of instruction (Stein, 2009), finding from the current study suggest these methods fail to produce the content knowledge gains of text-dependent instruction. Finding should not be interpreted as declaring other pedagogical forms such as museums, film, and group projects are holistically ineffective. Rather, analyses in the current study suggest that these forms of instruction do not provide optimal exposure if the instructional emphasis is content knowledge acquisition. While highly interactive and motivating, learning that accompanies these pedagogical models requires intellectual leaps between stimulating activities and content-specific knowledge (Stoddard, 2014). Traversing these gaps alone is challenging for most students and requires significant debriefing and discourse to process experiential learning in the context of existing schema. However, pushed for time and pressured for content coverage, teachers far too often forgo this process, leaving students with disjointed understanding.

The interaction effect between multimodal and text-dependent instruction further indicates that the former negates the positive content knowledge gains afforded by the latter. It is plausible that inverse achievement outcomes associated with multimodal instruction may be indicative of ineffective instructional uses. Wright-Maley (2015) found that teachers vacillate between "hard" and "soft" control over instruction, a tenuous task that leads to substantially variability in student learning. Likewise, Dack, van Hover, and Hicks (2015) noted how infrequently the use of experiential instruction effectively conveys historical/social studies content and concepts.

It is also important to recognize that NAEP-USH does not necessarily align with what multimodal instruction teaches. Perhaps this lack of association with knowledge attainment is an indicator that multimodal instruction develops different skills, like project-based learning (PBL), that, while motivating and effective for promoting affective dimensions of learning, do not readily translate well to test performance as measured by tests like the NAEP-USH. Moreover, the NAEP-USH did not include a quality indicator. Thus, the analysis was limited to frequency (or exposure) to multimodal instruction. Parker and colleagues (2013) found that students exposed to PBL scored higher than their counterparts on some measures of the AP Government exam, suggesting that quality instructional delivery is associated with student achievement. Nonetheless, when scaled together, multimodal exposure mitigated text-dependent effects. Faced with instructional choices and expectations to improve historical content knowledge, results suggest that text-dependent instruction is the optimal learning/pedagogical strategy given history knowledge as measured by NAEP-USH.

Careful Considerations of Accountability Policy and History Content Knowledge

It is no secret that few school leaders and even fewer teachers approve of the current wave of testing and accountability mandates pervasive in state educational policy across the nation. Moreover, the recent trend linking students' test scores to teacher performance is extremely controversial with research suggesting that many of the accountability models and associated value-added models failing to accurately align with recognized standards of good teaching (Amrein-Beardsley, 2008; Polikoff & Porter, 2014). The current study found that students in schools with a required high school history test correlated with average increased performance on the NAEP-USH compared to students in non-test buildings. This finding suggests a policy of curriculum standardization and accountability exposes students to test-taking skills and reinforces content knowledge acquisition. The presence of testing creates a continuum of access to history/social studies instruction not prevalent in the absence of testing.

While it is not the purpose of this research to condone the use of testing as a high-stakes component of teacher evaluation, the use of the test as a referee for curriculum suggests that testing or other state-level curricular mandates might help promote increased content knowledge among students. Given that the NAEP-USH 12th grade measures an aggregate of a students' US History content knowledge, results imply that students who receive greater exposure to US history are more likely to retain content-specific information. Curriculum standardization, and testing as a byproduct, can serve as a policy tool for guaranteeing content is taught and that all students have an opportunity to learn content. Given the differences found across schools, efforts to level curricular access, even in the form of accountability can provide greater content learning opportunities for students. However, similar to Marchant and colleagues (2006) NAEP analyses, we are hesitant to interpret our findings as decisively in favor of high-stakes test to support student learning in history. The effects of testing have to be weighed with the pedagogical costs placed upon teachers and students. We also recognize that there are substantial debates on what and how the history curricula and assessment should look and contend that this is a separate issue outside the scope of the current study. We encourage and welcome the continued debate on the purposes and practices of history and the social studies and caution against findings being viewed as essentialist in their recommendations or implications.

Limitations

Due to NAEP's sampling frame, researchers were unable to acquire representation from every state. Therefore, the study's use of state data as a school-level variable should be interpreted cautiously. The NAEP student survey is self-reported data; however, efforts were made to control

for response bias and effort by accounting for student interest in history as noted in previous research (Niemi & Junn, 1998). Moreover, research indicates that student reported NAEP data is more reliable than teacher reported data, particularly in later grades (Henke, Chen, & Goldman, 1999; Smith and Niemi, 2001). These limitations in the sampling frame prevented the development of a three-level multilevel model. The potential for between-state variance in students' historical content knowledge not examined in this model exists. In addition, student reported data of instructional exposure measured quantity, not quality. Interpretations regarding the efficacy of text-based instruction and rather ineffectiveness of multimodal instruction in this context should be interpreted cautiously. Lastly, critics argue that NAEP-USH is an imperfect tool for historical understanding. We agree the assessment does not represent the ceiling for historical understanding. However, it offers the largest and most representative sample of students' historical knowledge in the US and can serve as a baseline for determining, on a national scale, students' canonical content knowledge. Understanding students' knowledge of history has implications for how educators teach and drive policy for the discipline.

Conclusion

In this study, we examined students' history content knowledge as it related to sociocultural attributes, learning exposure, and school characteristics. Findings indicated that while student characteristics (such as race, gender, and socioeconomic status) were significantly associated NAEP-USH outcomes, students' exposure to instructional strategies remain substantial predictors of historical content knowledge. The researchers contend that while a sociocultural achievement gaps continue, issues related to curriculum, such as content, delivery, gender equity and access are important considerations for leveling historical knowledge. Pedagogically, text-dependent instruction illustrates the potential for disciplinary literacy as a high leverage practice for promoting historical knowledge and understanding. Interestingly, accountability mandates (via testing) may serve an important role in historical content knowledge development. Testing is associated with a continuum of access to history/social studies content, which has significant implications for all students. Lastly, the disproportional lack of achievement among students in predominately Black schools should give researchers and education stakeholders pause. However, current findings suggest that students are not destined to be fall within these gaps. Students of color, females, and less affluent students are not pre-ordained to perform poorly compared to wealthy, White/Asian, males. Increased efforts to expose students to disciplinary literacy skills coupled with supportive curricular policies that assure access to curriculum can positively affect students' learning of history.

References

Alexander, P. A. (2003). The development of expertise: The journey from acclimation to proficiency. *Educational Researcher*, 32(8), 10-14. <https://doi.org/10.3102/0013189X032008010>

Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education and Policy Analysis Archives*, 10(18). <https://doi.org/10.14507/epaa.v10n18.2002>

Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65-75. <https://doi.org/10.3102/0013189X08316420>

Amrein-Beardsley, A., & Berliner, D. C. (2003). Re-analysis of NAEP math and reading scores in states with and without high-stakes tests: Response to Rosenshine. *Educational Policy Analysis Archives*, 11(25). <https://doi.org/10.14507/epaa.v11n25.2003>

Au, W. (2007). High-stakes testing and curricular control: A qualitative metasynthesis. *Educational Researcher*, 36(5), 258-267. <https://doi.org/10.3102/0013189X07306523>

Ball, D. L. , & Forzani, F. M. (2011). Building a common core for learning to teach: And connecting professional learning to practice. *American Educator*, 35(2), 17.

Barton, K. C. (2012). History: Learning narratives to thinking historically. In W. B. Russell (Ed.), *Contemporary social studies: An essential reader* (pp. 119-138). Charlotte, NC: Information Age Publishing.

Barton, K. C., & Avery, P. G. (2016). Research on social studies education: Diverse students, settings, and methods. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of Research on Teaching* (fifth ed., pp. 985-1038). Washington, DC: American Educational Research Association. https://doi.org/10.3102/978-0-935302-48-6_16

Berliner, D. (2006). Our impoverished view of educational reform. *The Teachers College Record*, 108(6), 949-995. <https://doi.org/10.1111/j.1467-9620.2006.00682.x>

Borman, G. D. , & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research*, 78(3), 367-409. <https://doi.org/10.3102/0034654308321455>

Boyd, D., Lankford, H., Loeb, S., & Wyckoff, J. (2005). Explaining the short careers of high achieving teachers in schools with low performing students. *American Economic Review*, 95(2), 166-171. <https://doi.org/10.1257/000282805774669628>

Braun, H. (2004). Reconsidering the impact of high-stakes testing. *Education Policy Analysis Archives*, 12(1). <https://doi.org/10.14507/epaa.v12n1.2004>

Breakstone, J. (2014). Try, try, and try again: The process of designing new history assessments. *Theory and Research in Social Education*, 42, 453-485. <https://doi.org/10.1080/00933104.2014.965860>

Brophy, J. (2006). Graham Nuthall and social constructivist teaching: Research-based cautions and qualifications. *Teaching and Teacher Education*, 22, 529-537. <https://doi.org/10.1016/j.tate.2006.01.008>

Brown, K. D. (2011). Breaking the cycle of Sisyphus: Social education and the acquisition of critical sociocultural knowledge about race and racism in the United States. *The Social Studies*, 102, 249-255. <https://doi.org/10.1080/00377996.2011.563726>

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis*, 24(4), 305-331. <https://doi.org/10.3102/01623737024004305>

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology*, 23, 760-771. <https://doi.org/10.1002/acp.1507>

Chikkatur, A. (2013). Teaching and learning African American history in a multicultural classroom. *Theory and Research in Social Education*, 41(4), 514-534. <https://doi.org/10.1080/00933104.2013.838740>

Cho, S., & Reich, G. A. (2008). New immigrants, new challenges: High school social studies teachers and English language learners. *The Social Studies*, 99(6), 235-242. <https://doi.org/10.3200/TSSS.99.6.235-242>

Clotfelter, C. T. , Ladd, H. F. , & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4), 778-820. <https://doi.org/10.3368/jhr.XLI.4.778>

Common Core Anchor Standards (2010). Retrieved from: <http://www.corestandards.org/ELA-Literacy/CCRA/>

Cornbleth, C., & Waugh, D. (1995). *The great speckled bird: Multicultural politics and education policymaking*. Mahwah, NJ: Lawrence Erlbaum Associates.

Crocco, M. S. (2006). Gender and social education: What's the problem? In E. W. Ross (Ed.), *The social studies curriculum: Purposes and possibilities* (3rd ed., pp. 171-193). Albany: SUNY Press.

Crocco, M. S. (2008). Gender and sexuality in the social studies. In L. S. Levstik & C. A. Tyson (Eds.), *Handbook of research on social studies teaching and learning* (pp. 172-196). New York: Routledge.

Dack, H., van Hover, S., & Hicks, D. (2015). "Try not to giggle if you can help it": The implementation of experiential instructional techniques in social studies classrooms. *Journal of Social Studies Research*. <https://doi.org/10.1016/j.jssr.2015.04.002>.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, 97(2), 139-156. <https://doi.org/10.1037/0022-0663.97.2.139>

De La Paz, S., Felton, M., Monte-Sano, C., Croninger, R., Jackson, C., Deogracias, J. S. , & Hoffman, B. P. (2014). Developing historical reading and writing with adolescent readers: Effects on student learning. *Theory & Research in Social Education*, 42(2), 228-274. <http://dx.doi.org/10.1080/00933104.2014.908754>

Dee, T., & Jacob, B. A. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management*, 30(3), 418-446. <https://doi.org/10.1002/pam.20586>

Dorn, S. (2006). No more aggregate NAEP studies? [editorial]. *Education Policy Analysis Archives*, 14(31). <http://dx.doi.org/10.14507/epaa.v14n31.2006>

Education Week. (2010). Quality Counts.

Epstein, T. (2009). *Interpreting national history*. New York, NY: Routledge.

Executive summary: 50-state Report Card. (2010). *Education Week*, 29(17), 4. Retrieved from <http://www.edweek.org/ew/articles/2010/01/14/17execsum.h29.html?tkn=LRPFD2Notgkozr3C16sN%2FFO4NFx46gxyG4Qe&print=1>.

Fallace, T. (2010). Exploring three orientations to the social studies In E. E. Heilman (Ed.), *Social studies and diversity education: What we do and why we do it* (pp. 23-25). New York, NY: Routledge.

Fantuzzo, J. W., Lebeof, W. A., & Rouse, H. L. (2014). An investigation of the relations between school concentrations of student risk factors and student educational well-being. *Educational Researcher*, 43(1), 25-36. <https://doi.org/10.3102/0013189X13512673>

Fitchett, P. G. (2010). A profile of twenty-first century social studies teachers. *Journal of Social Studies Research*, 34(2), 229-265.

Fitchett, P. G., Heafner, T. L., & Lambert, R. G. (2014a). Assessment, autonomy, and elementary social studies time. *Teachers College Record*, 116, 1-33.

Fitchett, P. G., Heafner, T. L., & Lambert, R. (2014b). Examining social studies marginalization: A multilevel analysis. *Educational Policy*, 1-29. <https://doi.org/10.1177/0895904812453998>

Fogo, B. (2014). Core practices for teaching history: The results of a Delphi panel survey. *Theory & Research in Social Education*, 42(2), 151-196. <https://doi.org/10.1080/00933104.2014.902781>

Galston, W. A. (2007). Civic knowledge, civic education, and civic engagement: A summary of recent research. *International Journal of Public Administration*, 30, 623-642. <https://doi.org/10.1080/01900690701215888>

Gay, G. (2003). Deracialization in social studies teacher education textbooks. In G. Ladson-Billings (Ed.), *Critical race theory perspectives on the social studies: The profession, policies, and the curriculum* (pp. 123-147). Greenwich, CT: Information Age Press.

Giles, C., Wang, Y., Smith, J. & Johnson, D. (2013). "I'm no longer just teaching history." Professional development for teaching Common Core State Standards for literacy in social

studies. *Middle School Journal*, 44(3), 34-43. <https://doi.org/10.1080/00940771.2013.11461853>

Goldsmith, P. R. (2011). Coleman revisited: School segregation, peers, and frog ponds. *American Educational Research Journal*, 48(3), 508-535. <https://doi.org/10.3102/0002831210392019>

Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing in New York state on teaching social studies. *Teachers College Record*, 103(3), 398-426. <https://doi.org/10.1111/0161-4681.00120>

Guthrie, J. T., Klauda, S. L. , & Ho, A. N. (2012). Modeling the relationships among reading instruction, motivation, engagement, and achievement for adolescents. *Reading Research Quarterly*, 48(1), 9–26. <https://doi.org/10.1002/rrq.035>

Gutiérrez, R. (2008). A "Gap-Gazing" Fetish in Mathematics Education? Problematizing Research on the Achievement Gap. *Journal for Research in Mathematics Education*, 39(4), 357-364.

Hanushek, E. A., & Raymond, M. E. (2006). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, 24(2), 297-327. <https://doi.org/10.1002/pam.20091>

Heafner, T. L., & Fitchett, P. G. (2015). An opportunity to learn US history: What NAEP data suggest regarding the opportunity gap. *The High School Journal*, 98(3), 226-249. <https://doi.org/10.1353/hsj.2015.0006>

Henke, R. R., Chen, X., & Goldman, G. (1999). *What happens in classrooms? Instructional practices in elementary and secondary schools, 1994-95*. Washington, DC: US Department of Education, National Center for Educational Statistics.

Howard, T. (2003). The dis(g)race of the social studies. In G. Ladson-Billings (Ed.), *Critical race perspectives on social studies: The profession, policies, and curriculum* (pp. 27-44). Greenwich, CT: Information Age Publishing.

Ladson-Billings, G. (2003). Lies my teacher still tells: Developing a critical race perspective toward the social studies. In G. Ladson-Billings (Ed.), *Critical race theory perspectives on the social studies* (pp. 1-14). Greenwich, CT: Information Age Press.

Lampert, M. (2009). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education*, 61(1-2), 21-34. <https://doi.org/10.1177/0022487109347321>

Lazer, S. (2015). A large-scale assessment of historical knowledge and reasoning: NAEP U. S. history assessment. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 145-158). New York, NY: Routledge.

Lee, J. S., & Bowen, N. K. (2006). Parent involvement, cultural capital, and the achievement gap among elementary school children. *American Educational Research Journal*, 43(2), 193-218. <https://doi.org/10.3102/00028312043002193>

Lee, P. (1994). Historical knowledge and the national curriculum. In H. Bourdillon (Ed.), *Teaching History* (pp. 41-48). London, UK: Routledge.

Levinson, M. (2012). *No citizen left behind*. Cambridge, MA: Harvard University Press. <https://doi.org/10.4159/harvard.9780674065291>

Levstik, L. S. (2008). What happens in social studies classrooms? Research on K-12 social studies practice. In L. S. Levstik & C. A. Tyson (Eds.), *Handbook of Research in Social Studies Education* (pp. 50-64). New York, NY: Routledge.

Litner, T., & Schweder, W. (2008). Social studies in special education classrooms: A glimpse behind the door. *Journal of Social Studies Research*, 32(1), 3-9.

Marchant, G. J., Paulson, S. E., & Shunk, A. (2006). Relationships between high-stakes testing policies and student achievement after controlling for demographic factors in aggregated data. *Education Policy Analysis Archives*, 14(30). <https://doi.org/10.14507/epaa.v14n30.2006>

Maggioni, L., VanSledright, B., & Alexander, P. A. (2009). Walking on borders: A measure of epistemic cognition in history. *The Journal of Experimental Education*, 77(3), 187-214.
<https://doi.org/10.3200/JEXE.77.3.187-214>

Mickelson, R. A. (2001). Subverting Swann: First-and second-generation segregation in the Charlotte-Mecklenburg schools. *American Educational Research Journal*, 38(2), 215-252.
<https://doi.org/10.3102/00028312038002215>

Moje, E. B., & Speyer, J. (2008). The reality of challenging texts in high school science and social studies: How teachers can mediate comprehension. In K. Hinchman & H. Sheridan-Thomas (Eds.), *Best practices in adolescent literacy instruction* (pp. 185–211). New York: Guilford.

Monaghan, M. M. (2014). Gender, equity and education: Examining preservice teachers' perceptions. *Gender, Equity, Music, and Society*, 7(8), 4-12.

Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *Journal of the Learning Sciences*, 19(4), 539-568.
<https://doi.org/10.1080/10508406.2010.481014>

Monte-Sano, C. (2012). Building skills by doing history. *Phi Delta Kappan*, 94(3), 62-65.
<https://doi.org/10.1177/003172171209400314>

Monte-Sano, C., & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research*, 44(3), 273-299.
<https://doi.org/10.1177/1086296X12450445>

NAEP Assessment Governing Board, (2010). *U. S. history framework for the 2010 National Assessment of Educational Progress*. Washington, DC: US Department of Education.

National Center for Education Statistics. (2011). *The Nation's Report Card: US History 2010*. Washington, DC: Institute of Educational Sciences, Department of Education.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

Nichols, S. L., Glass, G. V., & Berliner, D. C. (2012). High-stakes testing and student achievement: Updated Analyses with NAEP data. *Educational Policy Analysis Archives*, 20(20), 1-35.
<https://doi.org/10.14507/epaa.v20n20.2012>

Niemi, R. G., & Junn, J. (1993). Civic courses and political knowledge of high school seniors. Paper presented at Annual Meeting of the American Political Science Association. Washington, DC.

Niemi, R. G., & Junn, J. (1998). *Civic education: What makes students learn*. New Haven, CT: Yale University Press.

Nokes, J. D. (2010). Observing literacy practices in history education classrooms. *Theory & Research in Social Education*, 38(4), 515–544. <https://doi.org/10.1080/00933104.2010.10473438>

Pace, J. (2011). The complex and unequal impact of high stakes accountability on untested social studies. *Theory & Research in Social Education*, 39(1), 32-60.
<https://doi.org/10.1080/00933104.2011.10473446>

Patterson, N. C. , Horner, S. L. , Chandler, P. , & Dahlgren, R. (2013). Who is at the gate? An examination of secondary social studies teachers support and curricular control in testing and non-testing states. In J. Pasie & P. G. Fitchett (Eds.), *The Status of Social Studies: Views from the Field* (pp. 289-300). Charlotte, NC: Information Age Publishing.

Parker, W. C. , Lo, J. , Yeo, A. J. , Valencia, S. W. , Diem, N. , Abbot, R. D. , . . . Vye, N. J. (2013). Beyond breadth-speed-test: Toward a deeper knowing and engagement in an advanced placement course. *American Educational Research Journal*, 50(6), 1424-1459.
<https://doi.org/10.3102/0002831213504237>

Podgursky, M. (2002). *NAEP background questions: What can we learn from NAEP about the effect of schools and teachers on student achievement?*. Washington, DC: National Assessment Governing Board.

Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teacher quality. *Educational Evaluation and Policy Analysis*, 1-18. <https://doi.org/10.3102/0162373714531851>.

Raudenbush, S., Bryk, A., Cheong, Y. F., Congdon, R., & du Toit, M. (2004). *HLM 6: Hierarchical linear and nonlinear modeling*. Lincolnwood, IL: Scientific Software International.

Reich, G. A. (2009). Testing historical knowledge: Standards, multiple-choice questions, and student reasoning. *Theory & Research in Social Education*, 37(3), 325-360. <https://doi.org/10.1080/00933104.2009.10473401>

Reich, G. A. (2011). Testing collective memory: Representing the soviet union on multiple-choice questions. *Journal of Curriculum Studies*, 43(4), 507-532. <https://doi.org/10.1080/00220272.2011.578665>

Reisman, A. (2012). Reading like a historian: A document-based history curriculum intervention in urban high schools. *Cognition and Instruction*, 30(1), 86-112. <https://doi.org/10.1080/07370008.2011.634081>

Reisman, A., & Wineburg, S. (2008). Teaching the skill of contextualization in history. *The Social Studies*, 99(5), 202-207. <https://doi.org/10.3200/TSSS.99.5.202-207>

Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal*, 50(1), 4-36. <https://doi.org/10.3102/0002831212463813>

Rosenshine, B. (2003). High-stakes testing: Another analysis. *Educational Policy Analysis Archives*, 11(24). <https://doi.org/10.14507/epaa.v11n24.2003>

Rothstein, R. (2004). We are not ready to assess history performance. *The Journal of American History*, 90(4), 1381-1391. <https://doi.org/10.2307/3660358>

Rumberger, R., & Palardy, G. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *The Teachers College Record*, 107(9), 1999-2045.

Salinas, C. (2006). Educating late arrival high school immigrant students: A call for a more democratic curriculum. *Multicultural Perspectives*, 8(1), 20-27. https://doi.org/10.1207/s15327892mcp0801_4

Saye, J., & The Social Studies Inquiry Research Collaborative. (2013). Authentic pedagogy: Its presence in social studies classrooms and relationship to student performance on state-mandated tests. *Theory & Research in Social Education*, 4(1), 89-132. <https://doi.org/10.1080/00933104.2013.756785>

Schmeichel, M. (2015). Skirting around critical feminist rationales for teaching women in social studies. *Theory & Research in Social Education*, 43(1), 1-27. <https://doi.org/10.1080/00933104.2014.1002648>

Schmidt, W. H. , Burroughs, N. A. , Zoido, P. & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher*, 44(7), 371-387. <https://doi.org/10.3102/0013189X15603982>

Segall, A. (2006). Teaching in the age of accountability. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 105-132). Greenwich, CT: Information Age Publishing.

Seixas, P., & Ercikan, K. (2015). Introduction: The new shape of history assessments. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 1-14). New York: Routledge.

Shanahan, C. (2009). Disciplinary comprehension. In S. Israel & G. Duffy (Eds.), *Handbook of Research on Reading Comprehension* (pp. 240-260). New York: Routledge.

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40-59.
<https://doi.org/10.17763/haer.78.1.v62444321p602101>

Shanahan, T., & Shanahan, C. (2012). What is disciplinary literacy and why does it matter? *Topics in Language Disorders*, 32(1), 7–18. <https://doi.org/10.1097/TLD.0b013e318244557a>

Smith, J., & Niemi, R. G. (2001). Learning history in school: The impact of course work and instructional practices on achievement. *Theory & Research in Social Education*, 29(1), 18-42.
<https://doi.org/10.1080/00933104.2001.10505928>

Smith, M., & Breakstone, J. (2015). Historical assessments of thinking: An investigation of cognitive validity. In K. Ercikan & P. Seixas (Eds.), *New directions in assessing historical thinking* (pp. 233-245). New York: Routledge.

Sousa, D. A. (2011). *How the brain learns*. Thousand Oaks, CA: Sage.

Stanford History Education Group (n. d.). *Reading like a historian*. Retrieved from: <http://sheg.stanford.edu/rh.html>.

Stein, P. (2009). Multimodal instructional practices. In J. Coiro, M. Knobel, C. Lankshear, & D. Leu (Eds.), *Handbook of Research on New Literacies* (pp. 871-898). New York: Routledge.

Stoddard, J. (2014). Teaching thoughtfully about and with film. *Social Education*, 78(5), 220-224.

Stodolsky, S. S., Salk, S., & Glaessner, B. (1991). Student views about learning math and social studies. *American Educational Research Journal*, 28(1), 89-116.
<https://doi.org/10.3102/00028312028001089>

Swanson, E., Wanzek, J., McCulley, L., Stillman-Spisak, S., Vaughn, S., Simmons, D., Fogarty, M., & Hairrell, A. (2015). Literacy and text reading in middle and high school social studies and English language arts classrooms. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 1-24. <https://doi.org/10.1080/10573569.2014.910718>.

Szpara, M. Y., & Ahmad, I. (2006). Making social studies meaningful for ELL students: Content and pedagogy in mainstream secondary classrooms. *Essays in Education*, 16.

Tanaka, A., & Murayama, K. (2014). Within-person analyses of situational interest and boredom: Interactions between task-specific perceptions and achievement goals. *Journal of Educational Psychology*, 106(4), 1122. <https://doi.org/10.1037/a0036659>

van Hover, S. D. (2006). Teaching in the Old Dominion. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 195-219). Greenwich, CT: Information Age Publishers.

VanSledright, B. (1998). On the importance of historical positionality to thinking about and teaching history. *International Journal of Social Education*, 12(2), 1-18.

VanSledright, B. (2008). Narrative of the nation-state, historical knowledge, and school history education. *Review of Educational Research*, 32, 109-146.
<https://doi.org/10.3102/0091732X07311065>

VanSledright, B. (2011). *The challenge of rethinking history education: On practices, theories, and policy*. New York, NY: Routledge.

VanSledright, B. (2014). *Assessing historical thinking and understanding*. New York: Routledge.

Volger, K. E. (2006). The impact high school graduation examination on Mississippi social studies teachers' instructional practices. In S. G. Grant (Ed.), *Measuring history: Cases of state-level testing across the United States* (pp. 273-302). Greenwich, CT: Information Age.

Volger, K. E. & Virtue, D. (2007). "Just the facts, ma'mam": Teaching social studies in the era of standards and high-stakes testing. *The Social Studies*, 98(2), 54-58.
<https://doi.org/10.3200/TSSS.98.2.54-58>

Wertsch, J. V. (2002). *Voices of collective remembering*. Cambridge, UK: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511613715>

Wineburg, S. (1998). Reading Abraham Lincoln: An expert/expert study in the interpretation of historical texts. *Cognitive Science*, 22(3), 319-346.
https://doi.org/10.1207/s15516709cog2203_3

Wineburg, S., Mosborg, S. , Porat, D. , & Duncan, A. (2007). Common belief and the cultural curriculum: An intergenerational study of historical consciousness. *American Educational Research Journal*, 44(1), 40-76. <https://doi.org/10.3102/0002831206298677>

Wineburg, S., & Reisman, A. (2015). Disciplinary literacy history: A toolkit for digital citizenship. *Journal of Adolescent & Adult Literacy*, 58(8), 636-639. <https://doi.org/10.1002/jaal.410>

Wright-Maley, C. (2015). On "stepping back and letting go": The role of control in the success or failure of social studies simulations. *Theory & Research in Social Education*, 43(2), 206-243.
<https://doi.org/10.1080/00933104.2015.1034394>

Zhang, T., Torney-Purta, J., & Barber, C. (2012). Students' conceptual knowledge and process skills in civic education: Identifying cognitive profiles and classroom correlates. *Theory and Research in Social Education*, 40(1), 1-34. <https://doi.org/10.1080/00933104.2012.649467>

Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement*, 26(1), 55-66. <https://doi.org/10.1111/j.1745-3984.1989.tb00318.x>

About the Authors

Paul G. Fitchett

University of North Carolina at Charlotte

Paul.Fitchett@uncc.edu

Paul Fitchett is Associate Professor of Teacher Education at the University of North Carolina at Charlotte Cato College of Education. In addition, he serves as the graduate director for the PhD in Curriculum and Instruction program. His research focuses on the intersections of teacher working conditions, education policy, and student learning outcomes with a particular focus on the social studies. His scholarship has been published in *Theory and Research in Social Education*, *Teachers College Record*, and *Education Policy*. He has also co-edited two books including the forthcoming title, *Social Studies in the New Educational Policy Era: Conversations on Purposes, Perspectives, and Practices*.

Tina L. Heafner

University of North Carolina at Charlotte

Tina.Heafner@uncc.edu

Tina Heafner is Professor of Teacher Education at the University of North Carolina at Charlotte Cato College of Education. Her administrative responsibilities include directing the College of Education Prospect for Success, M.Ed. in Secondary Education and the Minor in Secondary Education. Tina is the Vice-President-elect of the National Council for the Social Studies (NCSS) and serves as Past-Chair of the Executive Board of the NCSS College and University Faculty Council. Her research focuses on technology-mediated and online learning, content literacy and disciplinary inquiry, education policy and student learning outcomes with a particular focus on the social studies. Her scholarship has appeared in the *Journal of Technology and Teacher Education*, *High School Journal*, *Theory and Research in Social Education*, *Teachers College Record*, and *Education Policy*. She has also co-edited four books and co-authored eight books with her most recent title *Beginning Inquiry in US History: Short Texts for Inexperienced Readers*.

Richard G. Lambert

University of North Carolina at Charlotte

rglamber@uncc.edu

Richard Lambert is a Professor in the Department of Educational Leadership in the Cato College of Education at the University of North Carolina at Charlotte, Director of the *Center for Educational Measurement and Evaluation*, and Editor of *NHSA Dialog: A Research-to-Practice Journal for the Early Intervention Field*. His research interests include formative assessment for young children, applied statistics, and teacher stress and coping. His research has appeared in *Educational Policy Analysis Archives*, *Early Childhood Research Quarterly*, and the *Journal of Applied Measurement*, among other outlets.

education policy analysis archives

Volume 25 Number 65

June 26, 2017

ISSN 1068-2341



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or **EPAAP**. **EPAAP** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A1 (Brazil), SCImago Journal Rank; SCOPUS, SOCOLAR (China).

Please send errata notes to Audrey Amrein-Beardsley at audrey.beardsley@asu.edu

Join **EPAAP's Facebook community** at <https://www.facebook.com/EPAAPAAPE> and **Twitter feed** @epaa_aape.

education policy analysis archives
editorial board

Lead Editor: **Audrey Amrein-Beardsley** (Arizona State University)

Editor Consultant: **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **David Carlson, Margarita Jimenez-Silva, Eugene Judson, Mirka Koro-Ljungberg, Scott Marley, Jeanne M. Powers, Iveta Silova, Maria Teresa Tatto** (Arizona State University)

Cristina Alfaro San Diego State University	Ronald Glass University of California, Santa Cruz	R. Anthony Rolle University of Houston
Gary Anderson New York University	Jacob P. K. Gross University of Louisville	A. G. Rud Washington State University
Michael W. Apple University of Wisconsin, Madison	Eric M. Haas WestEd	Patricia Sánchez University of University of Texas, San Antonio
Jeff Bale OISE, University of Toronto, Canada	Julian Vasquez Heilig California State University, Sacramento	Janelle Scott University of California, Berkeley
Aaron Bevanot SUNY Albany	Kimberly Kappler Hewitt University of North Carolina Greensboro	Jack Schneider College of the Holy Cross
David C. Berliner Arizona State University	Aimee Howley Ohio University	Noah Sobe Loyola University
Henry Braun Boston College	Steve Klees University of Maryland	Nelly P. Stromquist University of Maryland
Casey Cobb University of Connecticut	Jaekyung Lee SUNY Buffalo	Benjamin Superfine University of Illinois, Chicago
Arnold Danzig San Jose State University	Jessica Nina Lester Indiana University	Sherman Dorn Arizona State University
Linda Darling-Hammond Stanford University	Amanda E. Lewis University of Illinois, Chicago	Adai Tefera Virginia Commonwealth University
Elizabeth H. DeBray University of Georgia	Chad R. Lochmiller Indiana University	Tina Trujillo University of California, Berkeley
Chad d'Entremont Rennie Center for Education Research & Policy	Christopher Lubienski University of Illinois, Urbana-Champaign	Federico R. Waitoller University of Illinois, Chicago
John Diamond University of Wisconsin, Madison	Sarah Lubienski University of Illinois, Urbana-Champaign	Larisa Warhol University of Connecticut
Matthew Di Carlo Albert Shanker Institute	William J. Mathis University of Colorado, Boulder	John Weathers University of Colorado, Colorado Springs
Michael J. Dumas University of California, Berkeley	Michele S. Moses University of Colorado, Boulder	Kevin Welner University of Colorado, Boulder
Kathy Escamilla University of Colorado, Boulder	Julianne Moss Deakin University, Australia	Terrence G. Wiley Center for Applied Linguistics
Melissa Lynn Freeman Adams State College	Sharon Nichols University of Texas, San Antonio	John Willinsky Stanford University
Rachael Gabriel University of Connecticut	Eric Parsons University of Missouri-Columbia	Jennifer R. Wolgemuth University of South Florida
Amy Garrett Dikkers University of North Carolina, Wilmington	Susan L. Robertson Bristol University, UK	Kyo Yamashiro Claremont Graduate University
Gene V Glass Arizona State University	Gloria M. Rodriguez University of California, Davis	

**archivos analíticos de políticas educativas
consejo editorial**

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editores Asociados: **Armando Alcántara Santuario** (Universidad Nacional Autónoma de México), **Jason Beech** (Universidad de San Andrés), **Angelica Buendia** (Metropolitan Autonomous University, Mexico), **Ezequiel Gomez Caride** (Pontificia Universidad Católica Argentina), **Antonio Luzon** (Universidad de Granada), **José Luis Ramírez** (Universidad de Sonora, Mexico)

Claudio Almonacid

Universidad Metropolitana de Ciencias de la Educación, Chile

Miguel Ángel Arias Ortega

Universidad Autónoma de la Ciudad de México

Xavier Besalú Costa

Universitat de Girona, España

Xavier Bonal Sarro Universidad

Autónoma de Barcelona, España

Antonio Bolívar Boitia

Universidad de Granada, España

José Joaquín Brunner Universidad

Diego Portales, Chile

Damián Canales Sánchez

Instituto Nacional para la Evaluación de la Educación, México

Gabriela de la Cruz Flores

Universidad Nacional Autónoma de México

Marco Antonio Delgado Fuentes

Universidad Iberoamericana, México

Inés Dussel, DIE-CINVESTAV,

México

Pedro Flores Crespo Universidad

Iberoamericana, México

Ana María García de Fanelli

Centro de Estudios de Estado y Sociedad (CEDES) CONICET, Argentina

Juan Carlos González Faraco

Universidad de Huelva, España

María Clemente Linuesa

Universidad de Salamanca, España

Jaume Martínez Bonafé

Universitat de València, España

Alejandro Márquez Jiménez

Instituto de Investigaciones sobre la Universidad y la Educación, UNAM, México

María Guadalupe Olivier Tellez,

Universidad Pedagógica Nacional, México

Miguel Pereyra Universidad de

Granada, España

Mónica Pini Universidad Nacional

de San Martín, Argentina

Miriam Rodríguez Vargas

Universidad Autónoma de Tamaulipas, México

José Gregorio Rodríguez

Universidad Nacional de Colombia, Colombia

Mario Rueda Beltrán Instituto de Investigaciones sobre la Universidad y la Educación, UNAM, México

José Luis San Fabián Maroto

Universidad de Oviedo, España

Jurjo Torres Santomé, Universidad de la Coruña, España

Yengny Marisol Silva Laya

Universidad Iberoamericana, México

Juan Carlos Tedesco Universidad

Nacional de San Martín, Argentina

Ernesto Treviño Ronzón

Universidad Veracruzana, México

Ernesto Treviño Villarreal

Universidad Diego Portales Santiago, Chile

Antoni Verger Planells

Universidad Autónoma de

Barcelona, España

Catalina Wainerman

Universidad de San Andrés, Argentina

Juan Carlos Yáñez Velasco

Universidad de Colima, México

**arquivos analíticos de políticas educativas
conselho editorial**

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editoras Associadas: **Geovana Mendonça Lunardi Mendes** (Universidade do Estado de Santa Catarina),
Marcia Pletsch, Sandra Regina Sales (Universidade Federal Rural do Rio de Janeiro)

Almerindo Afonso

Universidade do Minho
Portugal

Alexandre Fernandez Vaz

Universidade Federal de Santa
Catarina, Brasil

José Augusto Pacheco

Universidade do Minho, Portugal

Rosanna Maria Barros Sá

Universidade do Algarve
Portugal

Regina Célia Linhares Hostins

Universidade do Vale do Itajaí,
Brasil

Jane Paiva

Universidade do Estado do Rio de
Janeiro, Brasil

Maria Helena Bonilla

Universidade Federal da Bahia
Brasil

Alfredo Macedo Gomes

Universidade Federal de Pernambuco
Brasil

Paulo Alberto Santos Vieira

Universidade do Estado de Mato
Grosso, Brasil

Rosa Maria Bueno Fischer

Universidade Federal do Rio Grande
do Sul, Brasil

Jefferson Mainardes

Universidade Estadual de Ponta
Grossa, Brasil

Fabianny de Cássia Tavares Silva

Universidade Federal do Mato
Grosso do Sul, Brasil

Alice Casimiro Lopes

Universidade do Estado do Rio de
Janeiro, Brasil

Jader Janer Moreira Lopes

Universidade Federal Fluminense e
Universidade Federal de Juiz de Fora,
Brasil

António Teodoro

Universidade Lusófona
Portugal

Suzana Feldens Schwertner

Centro Universitário Univates
Brasil

Debora Nunes

Universidade Federal do Rio Grande
do Norte, Brasil

Lilian do Valle

Universidade do Estado do Rio de
Janeiro, Brasil

Flávia Miller Naethe Motta

Universidade Federal Rural do Rio de
Janeiro, Brasil

Alda Junqueira Marin

Pontifícia Universidade Católica de
São Paulo, Brasil

Alfredo Veiga-Neto

Universidade Federal do Rio Grande
do Sul, Brasil

Dalila Andrade Oliveira

Universidade Federal de Minas
Gerais, Brasil